DOCUMENT RESUME

ED 220 498 TM 820 511

AUTHOR Crooks, Terence J.
TITLE Generalization in Educational Research: Through a Glass Darkly.
PUB DATE Mar 82
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (66th, New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Educational Research; Evaluation Methods; *Generalization; Research Design; *Research Methodology; Research Problems; *Validity

ABSTRACT

A conceptualization of generalization and its place in educational research is proposed. Problems with existing approaches to generalization include limited concepts, misleading or simplistic results, perception variance or change, and the discouragement of sensitivity due to convenience. Researchers should distinguish between conclusions which are apparently specific to a study, and broader generalizations. Research should be conducted in natural settings, using methodology which interferes as little as possible with the characteristics of the setting. Researchers need to be meticulous in describing their sample, setting, treatments and instrumentation. To gain an adequate level of understanding, research efforts need to be more intensive and deeper than is customary. The choice of research focus (research questions, and hypotheses) is as important as methodology in carrying out the investigations. Extensive prior investigation is needed to identify key variables in the setting in which the research is to be carried out. Hypothesis testing should be de-emphasized, and the focus of research should be shifted to description, exploratory data analysis, estimation of the magnitude of effects, and replication of interesting findings. (Author/PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED220498

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

GENERALIZATION IN EDUCATIONAL RESEARCH: THROUGH A GLASS DARKLY

Terence J. Crooks
Higher Education Development Centre
University of Otago
New Zealand

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
T. J. Crooks
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

March, 1982

Paper presented at the Annual Meeting of the American Educational Research Association, New York City, March 19-23, 1982.

TM 820 511

INTRODUCTION

It has often been assumed or stated in treatises on educational research that the main or ultimate goal of such research is to reach conclusions which apply in a wide variety of educational settings and are essentially invariant over time. Such conclusions could then be regarded as "laws". In recent years, however, Lee Cronbach and others have argued, on both theoretical and empirical grounds, that it is neither practical nor productive to make broad generalizations (laws) the central goal of educational and psychological research.

In the next section of this paper, I review and extend the arguments against the current role which generalization plays in educational research, quoting from some of the most cogent writing on this topic. Then, in the following section, I sketch some proposals for the reform of educational research, proposals which would in my view lead to a more appropriate conception of generalization and its place in educational research.

PROBLEMS WITH EXISTING APPROACHES TO GENERALIZATION

Limited Concepts of Generalization

Researchers have too often adopted very limited concepts of generalization, in which almost all their attention was given to the issue of generalization from a sample of persons. This approach has been soundly criticised in much of the literature on research methodology. For example, Brunswik (1956), Hersen and Barlow (1976), Millman (1966) and Sidman (1960) all indicated that in addition to generalization across persons, researchers must consider generalization across situations and conditions.

Campbell (1957), in a pioneering article on internal and external validity of experiments, placed emphasis on generalization to variants of the treatment and of the observation (measurement) procedures. More recently, Cook and Campbell (1979) subdivided Campbell's previous concept of external validity into two aspects: construct validity of treatments and measures, and external validity. The former refers to the validity with which we can make generalizations about higher-order constructs from research operations, both manipulated independent variables and measured traits. As foreshadowed

in Campbell's 1957 paper, contruct validity can be assessed and enhanced by the use of multitrait-multimethod approaches, in a manner described in detail by Campbell and Fiske (1959). In its new, more restricted usage, external validity refers to the validity with which conclusions are drawn about generalizability to and across populations of persons, settings, and times.

Cronbach (1978, see also Cronbach, 1980) sums up my first point when he stresses that generalizations must take account of units, treatments, observing operations, and settings. He notes that:

> A study attends to certain units - individuals, perhaps, or classes. Second, there is a treatment - the curriculum materials, perhaps. A unit is exposed to a realization of that treatment. The teacher selects and organizes lessons, and proceeds in her own lucid or confusing style. Even with a supposedly standard treatment, realizations are certain to vary. Third, the evaluator obtains data before, during and/or after the treatment. He administers a certain form of test, or sends a certain visitor to the class to record certain kinds of impressions. These are the observing operations.
>
> When units are selected, local settings are selected incidentally. Studying teachers, we observe them in a social context, not as isolated individuals. The teacher is influenced by the supervision of administrative officers, and by any current tensions surrounding pay negotiations. Influences that differ from one locality to another are, in effect, confounded with the characteristics of the unit.
>
> There is a larger setting also. The intellectual and political climate of the period impinges on all units. Outcomes in one milieu or in one decade are not necessarily to be expected in a later decade. An evaluation of conscription in World War II, for example, does not indicate what a similar study would have found in 1969 [during the Vietnam war][pp. 202, 203].

Generalizations are Often Misleading or Simplistic

A second problem with current practices of generalization in educational research is that the generalizations produced are often misleading or at best simplistic.

In part, this problem arises from the sheer complexity of social interaction and cognitive functioning. As McGuire (1973) has noted:

> In our actual cognitive and social systems, effects are the outcome of multiple causes which are often in complex interactions; moreoever, it is the rule rather than the exception that the effects act back on the causal variables [p. 452]

Cronbach (1975), reviewing the findings of aptitude-treatment interaction research, commented as follows:

> Interactions are not confined to the first order; the dimensions of the situation and of the person enter into complex interaction [p. 116].
>
> Once we attend to interactions, we enter a hall of mirrors that extends to infinity. However far we carry out analysis - to the third or fifth order of any other - untested interactions of a still higher order can be envisioned [p. 119].
>
> This complexity forces us to ask once again, Should social science aspire to reduce behaviour to laws? [p. 116].

An even stronger view was expressed by Glass (1972):

> The laws of the social and behavioural sciences are of extremely limited generality. These laws are highly interactive with factors such as geography, time, culture, and characteristics of persons. If physical laws were as limited in generality as the laws so far discovered by social scientists, we would hesitantly creep out of bed each morning not knowing whether we would float to the ceiling or crash to the floor [p. 13].

Perhaps it would be helpful to briefly review two broad generalizations which have received much attention in educational psychology. The first is E.L. Thorndike's Law of Effect: that knowledge of results facilitates learning. McKeachie (1973), in an article entitled "The Decline and Fall of the Laws of Learning", reviewed the status of the Law of Effect and concluded that "the best generalization we can currently make [is]..... that knowledge of results eventuates in improved performance when the learner is motivated, when the knowledge of results is informative, and when the learner knows or is told what to do to correct his errors" (p. 10). This highly qualified generalization is a long way from the original Law of Effect. The second broad generalization I wish to examine is the claim that communicating behavioural objectives to students facilitates learning. Melton (1978), in a comprehensive review of the behavioural objectives literature, identified a number of conditions under which behavioural objectives appear to be ineffective:

> If students ignore the objectives provided, either because they are unaware of them, or because prior experience suggests that it is not important to take note of them.

> If the objectives are too general, or too ambiguous to be of particular assistance.
>
> If the objectives are of extreme facility or difficulty. (The structure of readability of instructional materials may be closely related to this condition).
>
> If the objectives of particular interest are only a small proportion of those provided to students.
>
> If students are so conscientious, or so highly motivated, that they achieve the objectives regardless of whether or not they are specified [p. 294].

Once again, the original broad generalization is seen to require considerable modification and qualification if it is to be in reasonable accord with empirical findings.

Another illustration of the dangers of generalization comes from research on the qualities which students desire in college teachers. Hundreds of studies have examined this issue, yet most have failed to even consider the possibility that the desired qualities might differ in different kinds of courses. For example, it seems likely that these qualities may be very different in advanced physics lectures than in introduction sociology discussion classes. To further complicate the situation, Feldman (1976) has demonstrated that the methodology by which students' preferences are assessed interacts with the findings of the research studies.

Another major reason why many generalizations are misleading is that educational research is often carried out in highly artificial ways or settings. In recent years, many researchers have stressed the need for greater attention to ecological validity. The writings of Egon Brunswik (1955, 1956) have been particularly influential. Petrinovich (1979) summarized Brunswik's concern as follows:

> Brunswik (1956) pointed out that systematic design almost inevitably involves the use of atypical backgrounds for the behaviours in question and embeds the behaviours in atypical contexts. This makes it difficult to detect the presence of interacting influences or to understand the magnitude of such influences, since many variables of potential importance have been removed experimentally. The systematic experimental method separates the variables controlling behaviour from the fabric in which they are embedded, and this destroys the pattern of correlations between variables as it exists in natural situations. The use of this approach has serious limiting consequences if the behavioural system under study is one that has multiple and intersubstitutable determiners, involves variable and multiple mediation, and is expressed in vicarious and hierarchical modes of response ..... If we wish to generalize the results of behavioural research to the contexts in which the behaviour normally occurs, we must not destroy the pattern of correlations that exists in such contexts [pp. 374, 375].

Brunswik advocated "representative" design, as an appropriate alternative to conventional "systematic" design. The idea of representative design, as it might be applied to research on teaching, has been rather well developed by Snow (1974), who noted that:

> The biggest threat to external validity may come when the experiment does not fit the nature of the behaviour being studied and, furthermore, does not include the means of discovering this fact [pp. 265, 266].
>
> Ecological validity concerns the extent to which the habitats or situations compared in an experiment are representative of the population of situations to which the investigator wishes to generalize [p. 272].

Another researcher who has made a major contribution to current thinking about ecological validity is Bronfenbrenner (1976), to whom we owe the classic statement that "much of developmental psychology is the science of the strange behavior of children in strange situations with strange adults for the briefest possible periods of time" (p.158). He has concluded that three requirements must be met if we are to make progress in the scientific study of educational systems and processes: our researches must for the most part be carried out in real-life educational settings; we must examine the relations between the characteristics of the learners and the several environments in which they live their lives (home, school, peer group, etc.), and between the different environments; and we must try to systematically contrast environmental systems or their structural components, as a means of investigating person-environment and environment-environment relations.

To sum up this section on the misleading or simplistic nature of many generalizations, I quote the conclusion reached by Cronbach and Snow (1977) in their massive work on aptitude-treatment interactions:

> It will be evident in what we have said throughout this book that we have serious reservations about the dominant style of research in instructional psychology. Insofar as educational reality is shaped by multiple parameters of student and treatment, traditional experiments cannot be powerful enough to support adequately complex conclusions. Insofar as variance arises out of the history of a particular class, it will be impossible to catch important causes, in the net of statistical inference [p. 519].

Generalizations change

Even when quite clear-cut and sound generalizations have been established, we cannot relax and enjoy the fruits of our labour. This is because, as Cronbach (1975) remarked, generalizations decay or change. We tend to speak of a scientific conclusion as if it were eternal, but there is much evidence that should encourage us to be more cautious. For instance, Cronbach noted that it was once a sound generalization that DDT kills mosquitos, until the evolutionary mechanism acted to make mosquitos resistant to DDT. Likewise, Bronfenbrenner (1958), in examining parenting practices in middle-class and lower-class families, found increasing permissiveness in both classes from 1930 to the 1950s, but differences between the classes in the 1950s were sometimes just the reverse of what had been observed in 1930. The great Russian psychologist Luria, summarizing his conclusions from the researches in Central Asia, stated that:

> The facts show convincingly that the structure of activity does not remain static during different stages of historical development and that the most important forms of cognitive processes - perception, generalization, deduction, reasoning, imagination, and analysis of one's own inner life - vary as the conditions of social life change and the rudiments of knowledge are mastered [Luria, 1976, p. 161].

In more speculative vein, Ghiselli (1974) has suggested that even such a well-established finding as the superiority of distributed practice over massed practice may not remain valid from one generation to another, and Cronbach (1980) has noted that:

> Perceptions of women and women's expectations of themselves have changed enough in recent years that older studies are an uncertain guide for today's personnel decisions. Today's woman may be staying in a job longer or shorter than the woman of a decade back. The old evidence that measures of nAch related to performance of men and not of women is obsolescent if not obsolete [p. 46].

In addition to difficulties resulting from changes in facts and relationships, we face the problem that educational research tends to lag behind contemporary educational practice. Glass (1972, p. 15) felt that published educational research had "an aroma of Chippendale about

it", and stated that "while the researcher is groping for understanding of the effects of televised instruction in large groups, the schools will have long since moved on to dual audio-television or video-cassettes imbedded in individualized instruction".

In view of these concerns, Cronbach (1975, p. 125) suggested that we are unable to store up generalizations and constructs for ultimate assembly into a network. He noted that social scientists have "modelled their work on physical science, aspiring to amass empirical generalizations, to restructure them into more general laws, and to weld scattered laws into coherent theory", but concluded that that lofty aspiration was far from realization and indeed that enduring systematic theories about people in society are not likely to be achieved.

While accepting the validity of some of Cronbach's (1975) comments, Weisz (1978) argued that the most specific, concrete, and circumstantial generalizations are the ones which are most likely to change with the passage of time, and that the most universal generalizations (those which are relatively general, abstract, and universalistic in form) are least likely to change with time. He felt that most of the examples of changing generalizations identified in the educational and psychological literature have been of the former type, and that Cronbach's pessimistic comments did not apply to the broader generalizations. However, it must be noted that the broadest and most universal generalizations are also those which are most susceptible to the arguments of the previous section: that the complexity of social phenomena is such that broad generalizations are often misleading or simplistic. Hence, the researcher seeking enduring generalizations appears to be caught between a researchers' Scylla and Charybdis.

Striving for Generalization may Discourage Sensitivity

A further concern about current approaches is that striving for broad generalization tends to discourage sensitivity in research. Variables and effects which are quite powerful in local settings or with individual subjects are disregarded or regarded merely as nuisances.

Indeed, variation in outcomes arising out of these local effects is requently treated as "error" when data are analysed. Because researchers are anxious to tie their findings to those of earlier studies, there is a tendency to use "established" variables and instruments uncritically, and not to seek to develop and include other, possibly context-specific, variables which may lead to new insights, and more powerful findings. Indeed, there is a form of inertia which leads to the use of instruments and procedures long after their deficiencies become evident.

A small study which I carried out (Crooks, 1980) illustrates this point rather well. Ten grade prediction variables were assessed in a large introductory physics course at college level. By far the most powerful predictor (r=.60) was the score on a short mathematics test designed to test the mathematical skills specifically required in the course. By contrast, the more usual predictor variables (ACT scores, SCAT scores, and high school rank) correlated no higher than 0.45 with final grades in physics. This study is unusual : the vast majority of grade prediction studies, including those aiming for "differential prediction", use only the widely available predictors. There are probably two main reasons for this : convenience, and the fact that the studies are aiming for rather broad generalizations.

This insensitivity to local variables and effects, arising out of a focus on abstract generalization, can extent even further, as McGuire (1973) has noted:

> In our holy determination to confront reality and put our theory to the test of nature, we have plunged through reality, like Alice through the mirror, into a never-never land in which we contemplate not life but data. All too often the scientific psychologist is observing not mind or behaviour but summed data and computer printout. He is thus a self-incarcerated prisoner in a platonic cave, where he has placed himself with his back to the outside world, watching its shadows on the wall. There may be a time to watch shadows, but not to the exclusion of the real thing [p. 453].

One particularly important aspect of this lack of sensitivity to the realities of interaction in local settings has been a neglect of careful observation procedures. As Cronbach (1975) put it:

> Originally, the psychologist saw his role as the scientific observation of human behaviour. When hypothesis testing became paramount, observation was neglected, and even actively discouraged by editorial policies of journals. Some authors now report nothing but F ratios [p. 124].

In similar vein, Sanders (1981, p.9) has noted that "in educational research, generally, we have skipped over the basic scientific step of observation of the facts pertaining to the phenomena we investigate".

Focussing on Generalization may Minimize Practical Impact

For a number of reasons, our dominant research model, with generalization as its central goal, has had comparatively little effect on educational practice. First, its findings have often (one could almost say usually) been equivocal, with conflicting results from different researchers in different settings. Second, local factors have often appeared to practitioners to invalidate a recommended approach or generalization. Third, researchers have often failed to address the major concerns or perceived needs of practitioners (Hogben, 1980). Fourth, the apparent requirements of scientific rigour in research have often led to strained and non-productive researcher-practitioner relationships (this point is very well developed by Runkel, 1975, pp.55, 56).

Stake (1978), in an article advocating more widespread use of case study approaches in social inquiry, has commented as follows on the relationship between research and its application:

> It is the legitimate aim of many scholarly studies to discover and validate laws. But the aim of the practical arts is to get things done. The better generalizations [for getting things done] often are those more parochial, more personal. In fields such as education and social work, where few laws have been validated and where inquiry can be directed toward gathering information that has use other than for the cultivation of laws, a persistent attention to laws is pedantic [p. 7].

Practitioners are likely to pay more attention if we focus on issues which concern them, using techniques which allow them to be more fully involved. If the local setting is treated as more than incidental, and real attention is paid to the forces at work in that situation, research is likely to be more meaningful and more susceptible to application. What is more, the interplay between researcher and practitioners which would occur under these circumstances is far more likely to lead to a productive consultant-practitioner relationship than in cases where the setting is seen merely as a source of "subjects".

TOWARD REFORM OF EDUCATIONAL RESEARCH

The following suggestions derive from the concerns expressed in the previous section. They represent an attempt to envisage a more productive plan for educational research than is evident in much existing practice. I make no claim for the originality of these ideas, but have found it useful to draw them together into a coherent grouping and attempt to elaborate them a little.

1. We need to recognise that broad generalization should be the goal of a field of research, not of a single empirical study. Far too commonly, we find researchers trying to reach very broad conclusions from very limited evidence. Not only does this frequently result in very inadequate or misleading generalizations, but it also diverts the researchers from paying adequate attention to the fine structure in their research setting and data. Stake (1978) quotes Francis Bacon as writing:

> There are and can be only two ways of searching and discovering truth. The one flies from the senses and particulars to the most general axioms ..... The other derives axioms from the senses and particulars, rising by a gradual and unbroken ascent, so that it arrives at the most general axioms last of all. This is the true way [p. 6].

A similar view was expressed by Cronbach (1975):

> Instead of making generalization the ruling consideration in our research, I suggest that we reverse our priorities. An observer collecting data in one particular situation is in a position to appraise a practice or proposition in that setting, observing effects in context. In trying to describe and account for what happened, he will give attention to whatever variables were controlled, but he will give equally careful attention to uncontrolled conditions, to personal characteristics, and to events that occurred during treatment and measurement. As he goes from situation to situation, his first task is to describe and interpret the effect anew in each locale, perhaps taking into account factors unique to that locale or series of events. As results accumulate, a person who seeks understanding will do his best to trace how the uncontrolled factors could have caused local departures from the modal effect. That is, generalization comes late, and the exception is taken as serious as the rule [pp. 124, 125].

The intensive focus on research synthesis over the past five years, following the 1976 article on meta-analysis by Gene Glass, has in my view been very worthwhile. It has suggested that there are "lowest common denominators" of knowledge accumulating through our empirical studies, and indeed some of these averaged effect-sizes have proved to be surprisingly large. However, it has also reminded

us of the very large unexplained variations from study to study, which indicate the need for considerable caution in applying any of the findings to a particular new situation. Epstein's work in personality (Epstein, 1979, 1980) carries a similar message : that there are general patterns in human behaviour, but that individual instances fluctuate wildly. Ken Boulding, among others, has noted that exceptional or unusual cases often lead to more useful insights and scientific advances than do the more average cases, and it seems clear that we will impede progress if we focus too early or too exclusively on broad generalization.

2. We need to distinguish, in our research reports, between conclusions which are apparently specific to a study, and broader generalizations, which are perhaps best viewed as predictions of the findings in future research. Thus most of our generalizations would best be viewed as somewhat refined hypotheses for future research, rather than anything more grandiose. Runkel (1978), speaking as a consultant on educational practice, believes

> We can make best use of experimentation if we treat the outcome as a possibly useful hypothesis for our next job of consultation, using the strengths of the study to strengthen our own imagination and the weaknesses to guide diagnosis [p. 72].

A related point is that we should frequently remind ourselves that generalization almost invariably involves a large component of human inference, in addition to any evidence from statistical inference.

3. Wherever feasible, our research should be conducted in natural settings, using methodology which interferes as little as possible with the characteristics of the setting. In recent years, there have been many advocates of "naturalistic" or "ethnographic" research approaches. Examples of such thinking are the writings of Willems and Rausch (1969) on naturalistic research, Stake (1978) on the use of case studies, and Parlett and Hamilton (1976) on "illuminative evaluation." Critics of these approaches have suggested that they tend to be "unscientific", but this is certainly not necessarily the case. Indeed, Cronbach and Suppes (1969) have observed that:

> Disciplined inquiry does not necessarily follow well established, formal procedures. Some of the most excellent inquiry is free-ranging and speculative in its initial stages, trying what might seem to be bizarre combinations of ideas and procedures, or restlessly casting about for ideas [p. 16].

However, there does seem to be a tendency among some of the advocates of naturalistic approaches to rely almost entirely on impressionistic data and the personal reactions of the researcher, and to shun quantitative methods. This amounts to throwing the baby out with the bathwater! The most naturalistic study will benefit from careful attention to triangulation of "measures", and from thorough documentation which allows the reader to assess the support for the researcher's conclusions. Equally, studies which are based on more obviously structured designs have much to gain from avoiding, wherever possible, distorting the natural setting in which the behaviour being studied takes place.

Scriven (1972) has made an important distinction in discussing the meanings of the terms "objective" and "subjective". He notes that subjective in the quantitative sense means based on the judgement of one individual, whereas subjective in the qualitative sense means unreliable, probably biased, and just a matter of opinion. Unfortunately, these two usages are often confused - the carefully reasoned and supported judgements of one individual are labelled "subjective" in the second sense of the term. Good naturalistic research may well be subjective in the quantitative sense, but will not be subjective in the qualitative sense.

4. A corollary of the previous point is that rigor and relevance are not incompatible or ipsative (Mehrens, 1978 ; Gibbs, 1979). Rigor without relevance is worthless, while relevance without rigor is of dubious value. Many who espouse rigor have been very careful and diligent about the internal validity of their research designs, but negligent about the external validity and construct validity of those same designs. This, together with the narrow focus of many of the studies, has led to a lack of relevance in educational research. We should aspire to be as rigorous in the pursuit of relevance as we are in the pursuit of internal validity.

5. We should make much greater use of the best in educational practice available to us. Glass (1972) has stated that

> There is probably far more knowledge (about how to manage and promote learning in the classroom) in the nervous systems of ten excellent teachers than an average teacher can distill from all the educational research journals in existence [p. 11]

> For the next several decades then, I am placing my money on evaluative inquiry which is integrated with the process of educational development in which master teachers play an important role [p. 17].

Likewise, Hawkins (1966), recognising that "the best practice exceeds the best theory in quite essential ways" (p. 3), argues that

> It seems to me the better part of wisdom to find the good school situations - not the better third but the best one per cent - and engage in close observation and intellectual resonance; then to try to recreate such situations and make them more abundant and reproducible, no holds barred [p. 4].

6. We should be meticulous in describing our sample, setting, treatments, and instrumentation. This enables interested readers to assess for themselves the applicability of our findings or generalizations to a different context, and to be more aware of the limitations of our findings. Unfortunately, this full detailing is discouraged by editorial policies for many journals, and also it is not always compatible with our desire to present our research in the best possible light.

7. Our research efforts need to be much more intensive and deep than is customary, if we are to gain an adequate level of understanding. The concept of "thick description", as described by Ryle (1971) and Geertz (1973), is an important one. When we are willing to consider more of the complexity of educational situations, shallow and unreliable generalizations will be discouraged. Central to the notion of thick description is the observation process. Without careful and intelligent observation, thick descriptions are not possible. A particularly valuable source of ideas about the process of intensive research is Cronbach's 1978 manuscript.

8. We need to recognise that our choice of research focuses (research questions, hypotheses) is at least as important as our methodology for carrying out the investigations. McGuire (1973) stresses this point very strongly:

> One drastic change that is called for in our teaching of research methodology is that we should emphasize the creative, hypothesis formation stage of research. It is my guess that at least 90% of the time in our current courses on methodology is devoted to presenting ways to test hypotheses and that little time is spent on the prior and more important process of how one creates these hypotheses in the first place [p.450].

Cronbach (1978) devotes an entire chapter to the process of identifying research questions. He identifies two phases of planning: The divergent phase of listing possible questions [appropriate to the setting and research orientation], and the convergent phase of assigning priorities among them. As he describes these states:

> The divergent phase identifies what could possibly be worth investigating. The convergent phase is dedicated to deciding what incompleteness is most acceptable [p. 419].

9. We need to be very wary about engaging in manipulative or even in correlational research without extensive prior investigation to identify key variables in the setting in which the research is to be carried out. Whitehead (1932), in his paper "The Rhythm of Education", identified three stages of mental growth : the stage of romance, the stage of precision, and the stage of generalization. Dockrell (1980) applied this three stage model to educational research, and indicated that both the stage of romance (naturalistic, exploratory research) and the stage of precision are necessary prerequisites to sound generalization. The progression of research approaches identified and adopted by Nuthall (1971) and his colleagues seems to me to be an example of excellent practice. Careful, extended naturalistic observation leads to identification of variables of potential interest. Methods of categorising and measuring those variables are then developed and validated, after which the more promising variables are included in wide-ranging correlational studies. Next, experiments are designed to manipulate variables and enable causal inference. These three steps are repeated to refine the findings, and ultimately explanatory theories may be developed. Such a progression of naturalistic, empirical, and theoretical research clearly is very time consuming, and demands a level of commitment which is very high. It would rarely be feasible for an isolated researcher, and indeed could keep a substantial research team busy for many years.

10. We need to de-emphasise hypothesis testing, and instead to focus on description, exploratory data analysis (Tukey, 1977), estimation of the magnitude of effects, and replication of interesting findings. The assumptions underlying hypothesis testing (such as random sampling and independence of data points) are often not met, and statistical significance is crucially dependent on sample size. Furthermore, statistical significance is widely misinterpreted

(Rozeboom, 1960: Carver, 1978). By focusing on statistical significance, we have frequently failed to consider the practical significance of the "significant" effects we observed, and on the other hand have been led to ignore non-significant but potentially meaningful findings. Cronbach (1975) suggests:

> Hereafter, let us see estimates of variance components and raw-score regression coefficient instead [of F-ratios], confidence intervals will serve adequately to keep us cautious. Let the author file descriptive information, at least in an archive, instead of reporting only those selected differences and correlations that are nominally "greater than chance". Descriptions encourage us to think constructively about results from quasi-replications, whereas the dichotomy significant/nonsignificant implies only a hopeless inconsistency..... There are more things in heaven or earth than are dreamt of in our hypotheses, and our observations should be open to them..... The observer should be a journalist, not a dramatist. To suppress a variation that might not recur is bad observing [p. 124].

11. When we do manipulate behaviour, using an experimental design, we should pay careful attention to the advice given by Snow (1974). He advocated the use of systematic replication (varying subjects, setting, and other design facets) to enhance generalization; intraexperiment observation to ensure the treatment is implemented and to assess the possible influence of extraneous events; extraexperiment observation to see the intervention in context and to judge longer-term outcomes and side effects; extensive description and analysis of characteristics of students and treatments; careful preparation of students for the experiment; and the use of lengthy treatments.

CONCLUSION

To sum up, I quote yet again from the writings of Lee Cronbach:

> Systematic inquiry can reasonably hope to make two contributions. One reasonable aspiration is to assess local events accurately, to improve short-run control. The other reasonable aspiration is to develop explanatory concepts, concepts that will help people to use their heads [Cronbach, 1975, p. 126].

As I see it, we have had a tendency to try to view the whole educational world "through a glass darkly", and it is time we spent more time and energy focussing on individual educational settings "face to face".

References

Bronfenbrenner, U. Socialization and social class through time and space. In E.E. Maccoby, T.M. Newcomb, & E.L. Hartley (Eds.), Readings in Social psychology (3rd ed). New York: Holt, Rinehart & Winston, 1958.

Bronfenbrenner, U. The experimental ecology of education. Teachers College Record, 1976, 78, 157-204.

Brunswik, E. Representative design and probabilistic theory in a functional psychology. Psychological Review, 1955, 62, 193-217.

Brunswik, E. Perception and the representative design of psychological experiments. Berkeley: University of California Press, 1956.

Campbell, D.T. Validity of experiments in social settings. Psychological Bulletin, 1957, 54, 297-312.

Campbell, D.T., & Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Carver, R.P. The case against statistical significance testing. Harvard Educational Review, 1978, 48, 378-399.

Cook, T.D., & Campbell, D.T. Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally, 1979.

Cronbach, L.J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-127.

Cronbach, L.J. Designing educational evaluations. Stanford University: Stanford Evaluation Consortium, 1978. (Preliminary version)

Cronbach, L.J. Selection theory for a political world. Public Personnel Management, 1980, 9, 37-50.

Cronbach, L.J., & Snow, R.E. Aptitudes and instructional methods: A handbook for research on interactions. New York: Irvington, 1977.

Cronbach, L.J., & Suppes. P. (Eds.) Research for tomorrow's schools: Disciplined inquiry for education. New York: MacMillan, 1969.

Crooks, T.J. Grade prediction: The usefulness of context-specific predictors. University of Otago, Dunedin, New Zealand, 1980. (ERIC Document Reproduction Service No. ED 194 547).

Dockrell, W.B. The contribution of research to knowledge and practice. In W.B. Dockrell & D. Hamilton (Eds.), Rethinking educational research. London: Hodder and Stoughton, 1980.

Epstein, S. The stability of behavior: I. On predicting most of the people much of the time. Journal of Personality and Social Psychology, 1979, 37, 1097-1126.

Epstein, S. The stability of behavior: II. Implications for psychological research. American Psychologist, 1980, 35, 790-806.

Feldman, K.A. The superior college teacher from the students' view. Research in Higher Education, 1976, 5, 243-288.

Geertz, C. The interpretation of cultures. New York: Basic Books, 1973.

Ghiselli, E. Some perspectives for industrial psychology. American Psychologist, 1974, 29, 80-87.

Gibbs, J.C. The meaning of ecologically oriented inquiry in contemporary psychology. American Psychologist, 1979, 34, 127-140.

Glass, G.V. The wisdom of scientific inquiry on education. Journal of Research in Science Teaching, 1972, 9, 3-18.

Glass, G.V. Primary, secondary, and meta-analysis of research. Educational Researcher, 1976, 5(10), 3-8.

Hawkins, D. Learning the unteachable. In L.S. Shulman & E.R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966.

Hersen, M., & Barlow, D.H. Single-case experimental designs: Strategies for studying behaviour change. New York: Pergamon Press, 1976.

Hogben, D. Research on teaching, and teacher training. The Australian Journal of Education, 1980, 24, 56-66.

Luria, A.R. Cognitive development: Its cultural and social fundations. Cambridge, Massachussetts: Harvard University Press, 1976.

McGuire, W.J. The Yin and Yang of progress in social pschology: Seven koan. Journal of Personality and Social Psychology, 1973, 26, 446-456.

McKeachie, W.J. The decline and fall of the laws of learning. Educational Researcher, 1974, 3(3), 7-11.

Mehrens, W.A. Rigor and reality in counseling research. Measurement and Evaluation in Guidance, 1978, 11, 8-13.

Melton, R.F. Resolution of conflicting claims concerning the effect of behavioural objective on student learning. Review of Educational Research, 1978, 48, 291-302.

Millman, J. In the service of generalization. Psychology in the Schools, 1966, 3, 333-338.

Nuthall, G.A. Teacher verbal behaviour and pupil learning. Unpublished manuscript, 1971. (Christchurch, New Zealand: University of Canterbury, Department of Education).

Parlett, M., & Hamilton, D. Evaluation as illumination: A new approach to the study of innovatory programs. In G.V. Glass (Ed.), Evaluation studies review annual (Vol. 1). London: Sage Publications, 1976.

Petrinovich, L. Probabalistic functionalism: A conception of research method. American Psychologist, 1979, 34, 373-390.

Rist, R.C. On the relations among educational research paradigms: From distain to detente. Anthropology and Education Quarterly, 1977, 8, 42-49.

Rozeboom, W.W. The fallacy of the null-hypothesis significance test. Psychological Bulletin, 1960, 57, 416-428.

Runkel, P.J. A consultant's view of research methods. In D. Bar-Tal & L. Saxe (Eds.), Social psychology of education: Theory and research. Washington, D.C.: Hemisphere Publishing Corporation, 1978.

Ryle, G. The thinking of thoughts: What is 'Le Penseur' doing? In G. Ryle Collected papers, London: Hutchinson, 1971.

Sanders, D.P. Educational inquiry as developmental research. Educational Researcher, 1981, 10(3), 8-13.

Scriven, M. Objectivity and subjectivity in educational research. In L.G. Thomas (Ed.), Philosophical redirection of educational research, Chicago: 71st yearbook of the National Society for the Study of Education (Part I), 1972.

Sidman, M. Tactics of scientific research. New York: Basic Books, 1960.

Snow, R.E. Representative and quasi-representative designs for research on teaching. Review of Educational Research, 1974, 44, 265-291.

Stake, R.E. The case study method in social inquiry. Educational Researcher, 1978, 7(2), 5-8.

Tukey, J.W. Exploratory data analysis. Reading, Mass.: Addison-Wesley, 1977.

Weisz, J.R. Transcontextual validity in development research. Child Development, 1978, 49, 1-12.

Whitehead, A.N. The aims of education. London: Ernest Benn, 1932.

Willems, E.P. & Rausch, H.L. (Eds.) Naturalistic viewpoints in psychological research. New York: Holt, Rinehart & Winston, 1969.